

## **PROTÓTIPO DE SISTEMA COMPUTACIONAL APOIADO POR INTELIGÊNCIA ARTIFICIAL: MAPEAMENTO AUTOMÁTICO DE FORMULÁRIOS MÉDICOS PARA BASES DE DADOS\***

ZALEWSKI W; LEE HD; SPOLAÔR N; MALETZKE AG; CAETANO AMJF; LORENA AC; COY CSR; FAGUNDES JJ; WU FC

Laboratório de Bioinformática/UNIOESTE/PTI e Serv. Coloproctologia/Faculdade de Ciências Médicas/UNICAMP

**OBJETIVOS:** Apresentar um protótipo computacional para o mapeamento automático de informações contidas em Formulários Médicos (FM) para Bases de Dados (BD) digitais, com auxílio de Inteligência Artificial (IA), por meio da simulação do mapeamento de FM definidos a partir de um protocolo para a doença de Crohn.

**MATERIAL E MÉTODOS:** O protótipo foi estruturado em quatro módulos: (1) geração de FM, (2) construção de padrões, (3) mapeamento de FM e (4) Reconhecimento de Caracteres Manuscritos (RCM). No módulo (1) são definidos atributos e regras para o preenchimento da BD e, após, é construído um modelo de FM composto por Campos de Múltipla Escolha (CME) e por Campos Numéricos (CN). Os CME permitem selecionar uma entre diversas opções, enquanto que os CN possibilitam a escrita de números. No módulo (2), a localização dos CME e dos CN no modelo de FM é identificada com o apoio do algoritmo de IA *K-Nearest Neighbor* e de histogramas de projeção vertical e horizontal. A identificação dos CN e CME preenchidos nos FM sob mapeamento é realizada no módulo (3). O módulo (4) possibilita o reconhecimento dos caracteres grafados nos CN por meio da extração de características e posterior indução de uma Rede Neuronal Artificial (RNA). O Laboratório de Bioinformática da UNIOESTE em parceria com o Serviço de Coloproctologia da UNICAMP elaboraram um protocolo referente à doença de Crohn, sobre o qual foi definido um subconjunto de informações utilizado na simulação do mapeamento automático de um modelo de FM composto por cinco CME e quatro CN. Foram distribuídas cinco cópias do modelo para dez colaboradores das áreas médica e computacional, totalizando 50 FM que foram preenchidos *ad libitum*, digitalizados e mapeados automaticamente para uma BD.

**RESULTADOS:** Para os Campos de Múltipla Escolha constatou-se uma precisão de 96,4%, e para o reconhecimento dos 504 caracteres grafados nos Campos Numéricos obteve-se uma precisão de 97,6%.

**CONCLUSÃO:** Os especialistas das áreas computacional e médica consideraram satisfatórios os resultados, visto que a precisão do mapeamento automático dos CME e dos CN presentes nos FM foi significativamente alta. Outro aspecto observado foi a sensível redução no tempo utilizado no mapeamento das informações descritas nos FM para BD. Em trabalhos futuros, serão avaliados outros algoritmos de IA e aplicadas outras técnicas de extração de características que possam aumentar a confiabilidade no RCM e a robustez do protótipo.

\*Trabalho apoiado pelo PDTA-FPTI/BR.